

特惠算力专区

产品介绍

文档版本 01
发布日期 2024-11-04



版权所有 © 华为云计算技术有限公司 2024。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为云计算技术有限公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为云计算技术有限公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为云计算技术有限公司

地址：贵州省贵安新区黔中大道交兴功路华为云数据中心 邮编：550029

网址：<https://www.huaweicloud.com/>

目录

1 什么是竞享实例?	1
2 竞享实例的优势.....	3
3 竞享实例的应用场景.....	4
3.1 图像与媒体.....	4
3.2 大数据分析.....	5
3.3 Web 应用.....	7
3.4 高性能计算.....	9

1 什么是竞享实例?

什么是竞享实例?

竞享实例是弹性云服务器（Elastic Cloud Server，ECS）竞价计费的一种新型售卖模式，其最核心的特点是超低价格和系统抢占机制，即您可以以较低折扣购买实例，但同时系统资源不足等极端情况下系统可能会自动回收这些实例。其适用于各种无状态、容错或者灵活的应用场景。

说明

【竞享实例购买入口】<https://activity.huaweicloud.com/ceci.html>

竞享实例的生命周期

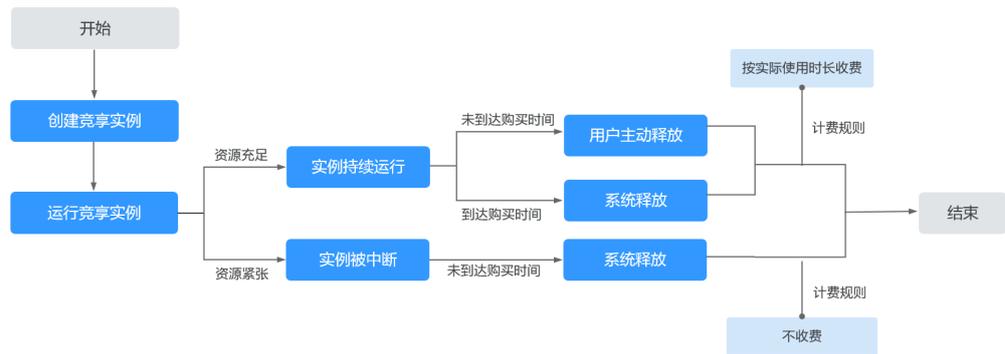
竞享实例创建时可选定购买时长（1-6小时），竞享实例创建后，使用时长达到用户购买时长时，实例将会进入“**中断回收中**”状态，随后释放。

竞享实例创建后，用户还可以根据业务需要随时主动释放实例，实例释放后按用户实际使用时长收费。

华为云会尽可能保障竞享实例正常运行，但在极端情况下会中断释放该类实例，实例被中断时会**秒级释放**，请在使用竞享实例时做好备份策略。购买时长为1-6个小时的实例被系统中断释放，可免除竞享实例的使用费用。

一般情况下，竞享实例回收会优先选择使用时长较短的实例中断回收，尽可能降低对用户的业务损伤，但仍然建议您优化应用设计，保证在竞享实例被中断回收的情况下业务也可以正常工作。您可以尝试主动释放竞享实例，检测在竞享实例被中断回收的情况下应用工作是否正常。

图 1-1 竞享实例的生命周期



说明

当前计费规则仅针对购买时长为1-6小时的竞享实例，购买时长超过6小时不在此计费规则内。

竞享实例的使用限制

- 华为云尽可能保障竞享实例正常运行，但在系统资源不足等极端情况下会主动释放该类实例
- 竞享实例目前只支持部分区域及规格，请以购买页面显示的资源为准
- 竞享实例不支持转换为按需和包周期实例
- 竞享实例不支持变更实例规格
- 竞享实例不支持变更操作系统
- 竞享实例不支持自动恢复功能

竞享实例的支持规格

竞享实例目前支持通用计算增强型、通用计算型、内存优化型、超高I/O型。

如线上暂无您需要的实例规格，可通过提工单申请，我们会根据实际情况，尽可能为您申请。

2 竞享实例的优势

- 性能保障
与按需实例使用相同底层ECS资源，与同类型同规格实例性能一致
- 成本降低
价格低至按需价格的1.5折，保障期内被系统回收可免除部分费用
- 灵活便捷
随租随用，秒速提升性能，支持设置时长，到期后自动释放
- 规模运行
支持超大规模工作负载，支撑企业快速完成批处理任务

说明

【竞享实例购买入口】<https://activity.huaweicloud.com/ceci.html>

3 竞享实例的应用场景

3.1 图像与媒体

离线渲染平台/影视动画渲染

- 场景概述
为了满足观众对于图形图像或影视特效要求不断提高，在渲染环节所占时间和投资比重也越来越大。传统的线下渲染方式已无法满足特效制作公司对于计算资源的重度需求，云化渲染逐渐成为主流。
- 客户瓶颈
以某部国产科幻电影的一个特效镜头为例，开场从地面拉向太空的2mins长镜头共有3000帧。以目前主流4K影片常用渲染配置——3000台24核服务器集群进行渲染，每台服务器并行渲染1帧，总时间需要20h，按需成本约32W。而在实际影视制作中，通常一个画面需要渲染，修改多次才可以出片，实际成本和时间会更高。
- 竞享实例的应用
相较于采用按需方式进行离线渲染，竞享实例可以渲染同样的特效画面，在时间相同的情况下只需8万元，成本下降75%。这些节省的成本还可以用来购买更多的竞享实例，实现成本和渲染周期的双重下降，并以此缩短整个电影拍摄制作周期，节省大量人力物力成本。

视频离线转码

- 场景概述
视频离线转码是指在离线状态下通过转码软件或转码平台，将音频、视频文件转码为适应各种终端（PC\TV\手机等）播放格式，主要应用于电视台、IPTV、OTT、有线网络、电信运营商、互联网视频等行业。
- 客户瓶颈
随着5G时代的到来，视频展现形式多样化，人们对视频质量要求提升。高清视频内容的制作与转换需求增加，对转码服务的效率和成本都是极大的挑战。已有资源不足以负荷转码任务量的增长，转码速度缓慢，如新增资源则是一笔不少的成本支出。
- 竞享实例的应用

客户可以将自己转码服务建立在竞享实例上，竞享实例价格优惠，可以以更低成本获取更高的性能。通过转码任务分发平台将视频任务切割成多个相互独立的计算任务，并分发到多台竞享实例中进行转码，转码完成后通过合成工具合并视频即可快速完成转码任务。

计算集群中CPU 总核数越多处理速度越快，竞享实例支持快速扩容使用，可灵活配置支撑转码任务高峰，结束任务后可以自由释放实例，减少资源浪费。利用竞享实例构建自己的转码服务，可以有效控制成本，提升转码速度。

📖 说明

【竞享实例购买入口】<https://activity.huaweicloud.com/ceci.html>

3.2 大数据分析

人工智能应用

- 场景概述

2016年AlphaGo横空出世，4:1战胜李世石，17年又以3:0战胜世界围棋冠军柯洁，此后三年，星际，Dota2，德州扑克等均涌现出超高水平AI。人工智能应用在其中起到了不可替代的作用。

游戏智能体通常采用深度强化学习方法，从0开始，通过与环境的交互和试错，学会观察世界、执行动作、合作与竞争策略。每个AI智能体是一个深度神经网络模型，主要包含如下步骤：

- a. 通过GPU分析场景特征（自己，视野内队友，敌人，小地图等）输入状态信息（Learner）。
- b. 根据策略模型输出预测的动作指令（Policy）。
- c. 通过CPU单线程模拟玩家，每个玩家（Actor）执行不同的策略。
- d. 不同策略对应不同结果，不同结果产生不同的Reward（奖励分数）。
- e. 该奖励分数作为参数用来更新策略模型，再进行新一轮学习。

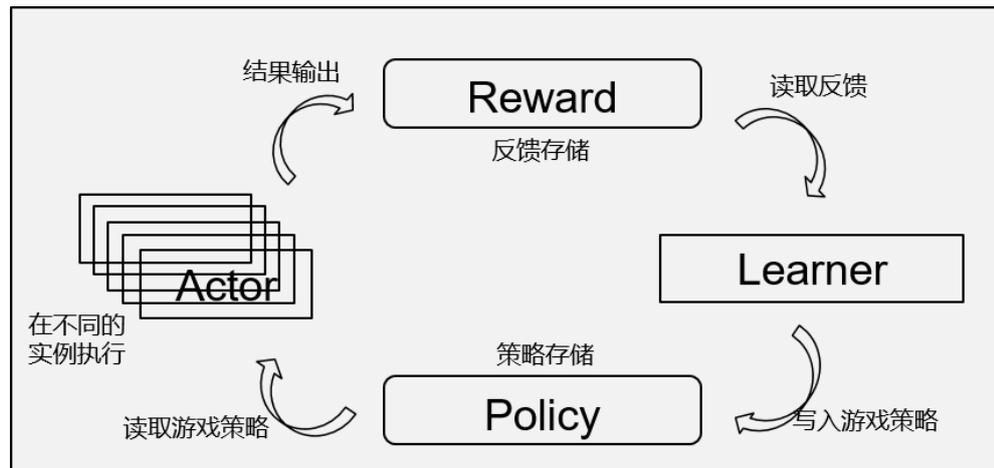
- 客户瓶颈

- a. 实时性与长期性：AI不仅要做出实时的操作决策，还要做出长期的规划决策，通常对于游戏时间30分钟左右的STG游戏，对应的决策步数（Policy）超过7000步，这意味着Actor执行Policy的时间成本较高。
- b. 复杂的动作空间：玩家需要同时操作移动方向、视角方向、攻击、姿态（站、蹲、趴、跳、跑）、交互（救人、拾取、换弹）等操作，产生复杂的组合动作空间，可行动作数量在 10^7 量级。对于CPU计算能力要求较高。
- c. 训练任务快速部署：客户进行AI强化学习时，需要短时间（10mins）拉起上万核CPU，对动态扩容能力要求较高。

- 竞享实例的应用

该AI学习引擎采用竞享实例提供CPU资源。得益于竞享实例的快速扩容与成本优势，引擎可以短时间生成超大规模AI（Actor）同时执行更多的策略，缩短模拟时间。而凭借竞享实例的强劲性能（全系C类型）该引擎训练一天相当于人类玩家打10万年。

图 3-1 人工智能应用架构图



Learner：学习集群，一般是多个GPU显卡组成训练集群

Actor：采用竞享实例提供CPU，每个线程作为一个AI玩家，用于测试策略的执行效果

Policy：Learner的输出结果，游戏AI的策略

Reward：Actor的执行结果的反馈，提供给Learner

重定向广告推广

- 场景概述

重定向广告（Retargeting）是一种基于应用、网页广告的定向技术，即针对广告受众（Audience）的浏览行为进行分析，在同一个广告位，推送为该用户定制的广告，实现千人千面。

- 客户瓶颈

重定向广告推广最终目的是将访问者转变为下单客户，是带动全球在线企业营收的主要因素之一。为了有效服务于广告，在极快的响应时间内实时中标，并通过自动化确保系统迅速响应竞价，广告推广商必须能够根据业务压力灵活快速并且成本可控地增加容量。

- 竞享实例的应用

客户使用包周期实例作为常规容量提供服务，在业务高峰时，得益于竞享实例低成本及快速扩缩容特性，竞享实例为系统提供可变容量以应对流量洪峰。自动化是这项业务的关键，所以客户需要进行业务容错性改造，实现任何一个或一些实例出现故障（被回收）时，可自行替换并继续运行，无需任何人工干预。

实时数据分析

- 场景概述

实时数据分析是指用适当的统计分析方法实时对收集来的大量数据进行分析，主要包含数据采集，加工，清洗，分析等环节。实时数据分析应用十分广泛，在车联网、金融保险、舆情分析、智慧城市等场景均有应用。

- 客户瓶颈

- a. 收集大数据时需要快速添加大量实例，并在收集结束后删除实例。
- b. 随着数据量不断增大，企业要高效准确地处理实时数据，需要运行越来越多的CPU资源来提供充足算力。采用按需实例会在成本可控上遇到较大挑战。

- 竞享实例的应用

客户通过使用竞享实例来降低用云成本，并在预算范围内尽可能的扩大集群规模，提升业务效率。客户要面对的最大挑战是一定概率的实例终止情况，通过保留一定量的按需实例作为竞享实例的BackUP可实现在竞享实例被回收时业务的平滑切换。

 说明

【竞享实例购买入口】 <https://activity.huaweicloud.com/ceci.html>

3.3 Web 应用

高弹性 Web 应用/无状态 Web 服务

- 场景概述

高弹性Web应用或无状态Web服务的访问量都有明显的波峰波谷变化，比如电商网站在大促时期，为确保Web应用服务稳定运行，需要配置更多服务器。

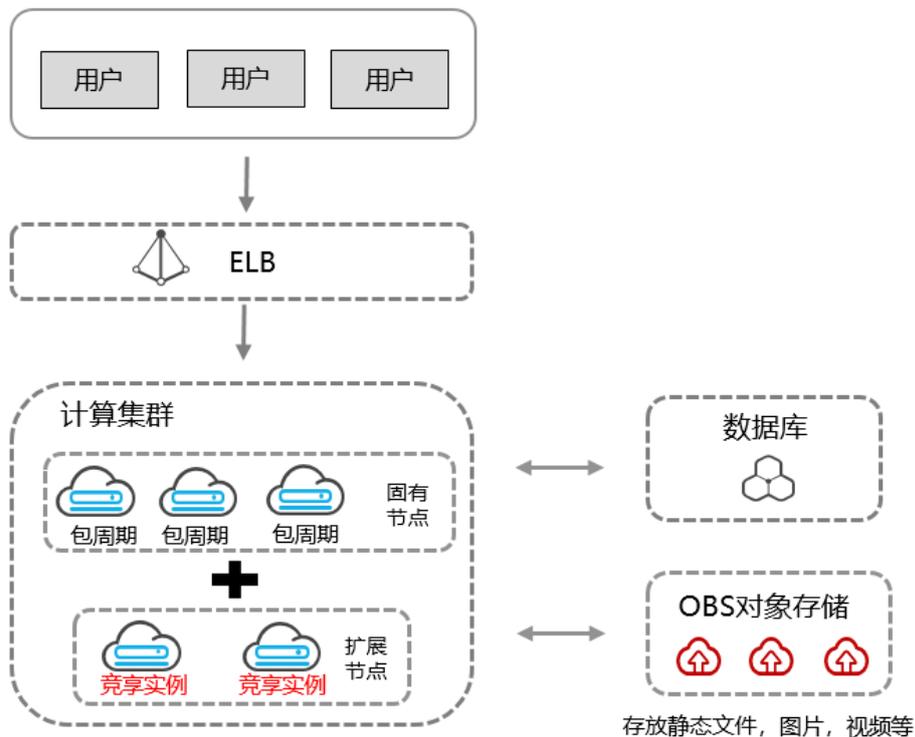
- 客户瓶颈

客户需要花费大量成本购置服务器，以维持业务高峰期的稳定运行。而高峰期时期较短，一般只持续短暂的几天甚至几小时。度过高峰期，访问量恢复时，配置过多的服务器会存在使用率不高等问题，造成了资源和成本的浪费。

- 竞享实例的应用

竞享实例具有随取随用的特点，是极具性价比的计算资源，用户可以在高峰期到来之前以极低的成本部署大量计算资源，构建临时计算集群，结合弹性负载均衡，将访问流量根据转发策略分发到后端多台临时弹性云服务器，以扩展应用系统对外的服务能力，平稳度过业务高峰期。高峰期过后访问量下降，对云服务器需求量降低时，用户可以释放掉多余的云服务器，节省用云成本。

图 3-2 高弹性 Web 应用/无状态 Web 服务场景架构



舆情获取业务

- 场景概述**

舆情获取业务是通过信息采集技术，对互联网信息进行快速有效的检索，以获取目标信息，主要应用于搜索引擎、新闻舆情获取、数据挖掘等。
- 客户瓶颈**

互联网具有海量页面信息，要对互联网中杂乱无章的信息进行处理，需要经过无数次的下载，分析，过滤、存储等过程，需要花费大量的时间，因此如何高效调用资源计算和分配任务才能提高处理效率至关重要。
- 竞享实例的应用**

为了高效采集页面数据，通常会多台服务器同时运作，并通过信息服务器统一管理，信息服务器会负责对全局URL队列进行分发，URL被发送至空闲云服务器上进行采集。服务器越多采集效率越高，可临时增加多台云服务器（竞享实例）以提高采集效率，同时，若任务量下降也可以减少竞享实例的使用，以节省成本。

CI/CD 与测试

- 场景概述**

CI/CD 是指使用软件交付管道以快速、自动化和可重复的方式从源代码生成发布版本的DevOps 工作流。其中包含了“持续集成”（CI）、“持续交付”（CD）、“持续部署”、“持续测试”等多个部分。
- 客户瓶颈**

CI/CD 对构建时效性要求高，需要快速在预生产环境完成业务部署及测试，若一切正常则会快速上线生产环境。程序员希望在提交代码后尽快返回结果，但业务

对资源的占用存在明显的波峰波谷，高峰期代码提交集中，计算资源紧张导致任务堵塞，耗时长。波谷时期大量资源被闲置造成资源浪费。

- 竞享实例的应用

我们可以借助多种插件和管理工具提升CI/CD系统的弹性扩容能力，在构建任务的时候，判断现有资源池是否足够负荷，若超出负荷范围则自动为资源池增加新的竞享实例，提高运算能力，当任务结束，可以自动把这些竞享实例释放掉，以节省成本。

3.4 高性能计算

生物医药（基因测序）

- 场景概述

生物医药包含基因测序，DNA变异检测等技术，随着人们对于健康关注度不断提高，基因测序等现代健康检测和预测手段正快速推广普及。基因测序是通过测序仪将不可视的生物信息转换为可视的文本信息，并依托计算、存储和大数据等资源，结合生物信息算法对文本信息进行分析和研究。

- 客户瓶颈

基因测序机构随着其业务的不断发展，现有存储系统遇到了巨大的挑战，每日需对数万个基因组进行统计分析，基因组数据所产生的数据容量多达几十TB，数据计算量激增，导致分析研究周期变长，基因测序机构急需投入成本和资源以提升计算速度。

- 竞享实例的应用

竞享实例提供了多种规格资源，并且成本极低，基因测序机构可以根据业务需要选择适用的规格，并依赖管理调度平台，弹性扩展自有的高性能计算集群规模，在业务量高峰时段，可快速提升机算速度，缩短计算周期，节省大量硬件设施投入，优化投入产出成本。

能源勘探

- 场景概述

为了满足全球日益增长的能源需求，石油和天然气行业已经成为高性能计算（HPC）的最大应用和创新行业之一。随着自然资源消耗，以及勘探成本增加，高度复杂的计算建模方式取代了传统的经验+实验方法，成为能源勘探和开发的重要工具。

- 客户瓶颈

能源公司非常依赖于地震数据，以及从数据中获得的模型，从而为勘探过程提供洞察和指导。地震数据的高质量 and 快速处理，分析对一个成功的勘探和项目开发起到至关重要的作用。一项地震调查可能会带来数百TB的数据，即使在超级计算机也要花费数月的时间来处理。总体上这类场景总是受到时间，计算硬件和软件的限制。

- 竞享实例的应用

根据Rock Flow Dynamics统计，一项石油勘探工作采用本地计算集群进行分析，需要2-8年时间来完成。如果采用云化+竞享实例方式，该时间可以缩短至数周。并且可以节约一大笔硬件采购费用以及后续维护费。使用竞享实例配合OBS等数据存储服务可以创建经济高效，可扩展的计算集群，为能源勘探解决方案提供高性价比算力支持。

芯片设计

- 场景概述

芯片设计业处于半导体行业的最上游，随着半导体制程工艺的逐步演进，芯片设计行业迅猛发展。现代芯片设计则是通过EDA等电子设计自动化工具来辅助设计人员完成工作。

- 客户瓶颈

芯片设计作为高性能计算的应用场景之一，对于IT基础设施的性能要求十分严苛，单一计算任务调度成百上千核的CPU。芯片设计企业的工作负载通常具有较强的周期性，不论是因多个项目并行工作造成的短期高峰，还是单个项目排期导致的长周高峰，企业都必须花费高额费用配置设备资源以满足高峰期需求，但实际年化使用率不足10%，造成资源浪费。

- 竞享实例的应用

芯片设计企业可以通过部署固定数量的包周期云服务器资源以维持日常负载需求，并采用竞享实例作为短期资源部署，以应对业务周期性的增长而带来的突发负载，优化资源配比以获取高性价比计算资源，为企业提高效能，节省运营成本。

说明

【竞享实例购买入口】<https://activity.huaweicloud.com/ceci.html>